

Klasifikasi multikelas kanker otak dengan metode multiple support vector machine recursive feature elimination dan twin support vector machine = Multiclass brain cancer classification using multiple support vector machine recursive feature elimination and twin support vector machine

Vinezha Panca, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20444707&lokasi=lokal>

Abstrak

ABSTRAK

Kanker merupakan salah satu penyebab kematian terbesar di seluruh dunia. Secara khusus, kanker otak adalah kanker yang terjadi pada sistem saraf pusat. Salah satu hal yang dapat dilakukan untuk penelitian kanker otak menggunakan machine learning adalah melakukan pendeteksian jenis kanker otak dengan memanfaatkan microarray data. Permasalahan tersebut merupakan masalah klasifikasi multikelas. Dengan menggunakan pendekatan one versus one, akan terbentuk sebanyak $k-1/2$ masalah dua kelas, di mana k menunjukkan jumlah kelas. Karena data kanker otak memiliki fitur yang sangat banyak, perlu dilakukan seleksi fitur. Pada penelitian ini, akan diimplementasikan metode Multiple Multiclass Support Vector Machine Recursive Feature Elimination MMSVM-RFE sebagai metode seleksi fitur, dan Twin Support Vector Machine TWSVM sebagai metode klasifikasi. Pada metode MMSVM-RFE dilakukan pelatihan SVM-RFE pada setiap masalah dua kelas, sehingga setiap masalah dua kelas memiliki pengurutan fitur masing-masing. Sebagai metode klasifikasi, TWSVM memiliki tujuan untuk mencari hyperplane masing-masing kelas sedemikian sehingga data kelas satu sedekat mungkin terhadap suatu hyperplane namun sejauh mungkin dengan hyperplane lainnya. Rata-rata akurasi tertinggi pada simulasi menggunakan kernel linear pada MMSVM-RFE dan kernel linear pada TWSVM adalah 95,33 dengan menggunakan 200 fitur. Rata-rata akurasi tertinggi pada simulasi menggunakan kernel linear pada MMSVM-RFE dan kernel RBF pada TWSVM adalah 87 dengan 70 fitur. Sedangkan apabila proses validasi juga dilakukan pada seleksi fitur, rata-rata akurasi tertinggi yang diperoleh adalah 90,67 dengan menggunakan 90 fitur.

<hr>

ABSTRACT

Cancer is one of main causes of death worldwide. Brain cancer is a type of cancer which occurs at central nervous system. Taking advantage from microarray data, machine learning methods can be applied to help brain cancer prediction according to its types. This problem can be referred as a multiclass classification problem. Using one versus one approach, the multiclass problem with k classes can be transformed into $k-1/2$ binary class problems. The huge amount of features makes it necessary to use feature selection. In this research, Multiple Multiclass Support Vector Machine Recursive Feature Elimination MMSVM RFE method is implemented as the feature selection method, and Twin Support Vector Machine TWSVM method is implemented as the classification method. The main concept of MMSVM RFE is to train SVM RFE at each binary problem so that each binary problem will have their own arrangements of feature. As a classification method, TWSVM is trained to find two hyperplanes, each representative of its own class. The data of one class must be as near as possible from its representative hyperplane while also must be as far as possible from the other hyperplane. In the simulation which uses linear kernel on MMSVM RFE and linear

kernel on TWSVM, the highest average accuracy is 95,33 , using 200 features. In the simulation which uses linear kernel on MMSVM RFE and RBF kernel on TWSVM, the highest average accuracy is 87 , using 70 features. In the case where the feature selection process is included in doing validation, the highest average accuracy is 90,67 , using 90 features.